

#4

35.C15311



PATENT APPLICATION

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:)	
	:	Examiner: Not Assigned
Fumiaki ITO, et al.)	
	:	Group Art Unit: 2641
Application No.: 09/839,309)	
	:	
Filed: April 23, 2001)	
	:	
For: VOICE BROWSER APPARATUS)	August 20, 2001
AND VOICE BROWSING METHOD	:	

Commissioner for Patents
Washington, D.C. 20231

CLAIM TO PRIORITY

Sir:

Applicants hereby claim priority under the International Convention and all rights to which they are entitled under 35 U.S.C. § 119 based upon the following Japanese

Priority Application:

JAPAN

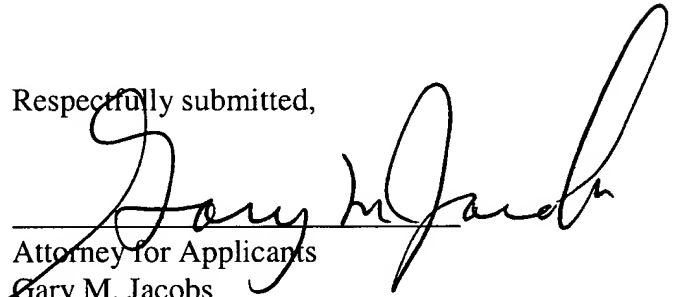
2000-127525

April 27, 2000.

A certified copy of the priority document is enclosed.

Applicants' undersigned attorney may be reached in our Washington, D.C.
office by telephone at (202) 530-1010. All correspondence should continue to be directed to our
address given below.

Respectfully submitted,

A handwritten signature in black ink, appearing to read "Gary M. Jacobs", is written over a horizontal line.

Attorney for Applicants

Gary M. Jacobs

Registration No. 28,861

FITZPATRICK, CELLA, HARPER & SCINTO
30 Rockefeller Plaza
New York, New York 10112-3801
Facsimile: (212) 218-2200

GMJ/cmv

CEA 15311 US /
04/ 839, 309
Fumiyaki MD, ETAL.
4-23-07



別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出願年月日
Date of Application:

2000年 4月27日

出願番号
Application Number:

特願2000-127525

出願人
Applicant(s):

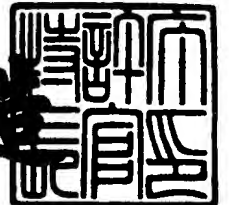
キヤノン株式会社

CERTIFIED COPY OF
PRIORITY DOCUMENT

2001年 5月18日

特許庁長官
Commissioner,
Japan Patent Office

及川耕造



【書類名】 特許願

【整理番号】 4054023

【提出日】 平成12年 4月27日

【あて先】 特許庁長官 近藤 隆彦 殿

【国際特許分類】 G06F 17/21

【発明の名称】 文書処理装置及びその方法、及びそのプログラムを格納した記憶媒体

【請求項の数】 14

【発明者】

 【住所又は居所】 東京都大田区下丸子3丁目30番2号キャノン株式会社
内

 【氏名】 伊藤 史朗

【発明者】

 【住所又は居所】 東京都大田区下丸子3丁目30番2号キャノン株式会社
内

 【氏名】 藤井 憲一

【発明者】

 【住所又は居所】 東京都大田区下丸子3丁目30番2号キャノン株式会社
内

 【氏名】 上田 隆也

【発明者】

 【住所又は居所】 東京都大田区下丸子3丁目30番2号キャノン株式会社
内

 【氏名】 池田 裕治

【特許出願人】

 【識別番号】 000001007

 【住所又は居所】 東京都大田区下丸子3丁目30番2号

 【氏名又は名称】 キャノン株式会社

 【代表者】 御手洗 富士夫

【電話番号】 03-3758-2111

【代理人】

【識別番号】 100090538

【住所又は居所】 東京都大田区下丸子3丁目30番2号キャノン株式会社
内

【弁理士】

【氏名又は名称】 西山 恵三

【電話番号】 03-3758-2111

【選任した代理人】

【識別番号】 100096965

【住所又は居所】 東京都大田区下丸子3丁目30番2号キャノン株式会
社内

【弁理士】

【氏名又は名称】 内尾 裕一

【電話番号】 03-3758-2111

【選任した代理人】

【識別番号】 100110009

【住所又は居所】 東京都大田区下丸子3丁目30番2号キャノン株式会
社内

【弁理士】

【氏名又は名称】 青木 康

【電話番号】 03-3758-2111

【手数料の表示】

【予納台帳番号】 011224

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

特 2 0 0 0 - 1 2 7 5 2 5

【包括委任状番号】 9908388

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 文書処理装置及びその方法、及びそのプログラムを格納した記憶媒体

【特許請求の範囲】

【請求項 1】 指定された取得先より所定のマークアップ言語で記述された文書を取得する文書取得手段と、

音声入出力内容を規定するルールを所定の複数のルールより選択するルール選択手段と、

前記ルール選択手段で選択されたルールに基づいて、前記文書取得手段で取得された文書の指定範囲を解析して、音声出力内容と、音声入力候補と、各音声入力候補に対応する次の処理対象を指定する指定情報とを取り出す文書解析手段と、

前記文書解析手段で取り出された音声出力内容を音声出力する音声出力手段と、

ユーザからの音声入力を音声認識する音声認識手段と、

前記音声認識手段による認識結果を、前記文書解析手段で取り出された入力候補と照合して、当該認識結果に一致する入力候補に対応する指定情報に基づいて、前記文書取得手段による新たな文書の取得あるいは前記文書解析手段による次の解析を制御する制御手段とを備えることを特徴とする文書処理装置。

【請求項 2】 前記ルール選択手段が、前記文書取得手段で取得された文書中に記述されたルール識別情報に基づいて、ルールを選択することを特徴とする請求項 1 記載の文書処理装置。

【請求項 3】 前記ルール識別情報は所定のタグの所定の属性値であることを特徴とする請求項 2 記載の文書処理装置。

【請求項 4】 前記取得された文書中にルール識別情報が記述されていない場合には、前記ルール選択手段が所定のルールを選択することを特徴とする請求項 2 記載の文書処理装置。

【請求項 5】 前記文書解析手段が、前記指定情報として次の文書の取得先を取り出すことを特徴とする請求項 1 記載の文書処理装置。

【請求項 6】 前記文書解析手段が、前記指定情報として次の文書の解析範囲を取り出すことを特徴とする請求項 1 記載の文書処理装置。

【請求項 7】 前記ルール選択手段が、ユーザの指示に基づいて、ルールを選択することを特徴とする請求項 1 記載の文書処理装置。

【請求項 8】 前記ルール選択手段が、前記ユーザの指示に基づくルールと、前記文書取得手段で取得された文書中に記述されたルール識別情報に基づくルールとの中から、所定の一方を優先させてルールを選択することを特徴とする請求項 7 記載の文書処理装置。

【請求項 9】 前記複数のルールが、前記文書中で所定のタグの所定の属性値を音声出力内容とし、第 2 の所定のタグで囲まれた内容をそれぞれ入力候補とするルールを含むことを特徴とする請求項 1 記載の文書処理装置。

【請求項 10】 前記ルールでは、前記認識結果がある入力候補と一致した場合に、その入力候補に対応する前記第 2 の所定のタグで囲まれた内容以降の第 3 の所定のタグまでを次の音声出力内容とし、当該音声出力内容中のアンカーを次の入力候補とすることを特徴とする請求項 9 記載の文書処理装置。

【請求項 11】 前記複数のルールが、前記文書の先頭から所定のタグまでを音声出力内容とし、当該音声出力内容中のアンカーを入力候補とするルールを含むことを特徴とする請求項 1 記載の文書処理装置。

【請求項 12】 前記音声入力及び音声出力が、電話回線を介して行なわれることを特徴とする請求項 1 記載の文書処理装置。

【請求項 13】 指定された取得先より所定のマークアップ言語で記述された文書を取得する文書取得工程と、

音声入出力内容を規定するルールを所定の複数のルールより選択するルール選択工程と、

前記ルール選択工程で選択されたルールに基づいて、前記文書取得工程で取得された文書の指定範囲を解析して、音声出力内容と、音声入力候補と、各音声入力候補に対応する次の処理対象を指定する指定情報とを取り出す文書解析工程と

前記文書解析工程で取り出された音声出力内容を音声出力する音声出力工程と

ユーザからの音声入力を音声認識する音声認識工程と、

前記音声認識工程による認識結果を、前記文書解析工程で取り出された入力候補と照合して、当該認識結果に一致する入力候補に対応する指定情報に基づいて、前記文書取得工程による新たな文書の取得あるいは前記文書解析工程による次の解析を制御する制御工程とを備えることを特徴とする文書処理方法。

【請求項 1 4】 指定された取得先より所定のマークアップ言語で記述された文書を取得する文書取得工程と、

音声入出力内容を規定するルールを所定の複数のルールより選択するルール選択工程と、

前記ルール選択工程で選択されたルールに基づいて、前記文書取得工程で取得された文書の指定範囲を解析して、音声出力内容と、音声入力候補と、各音声入力候補に対応する次の処理対象を指定する指定情報とを取り出す文書解析工程と

前記文書解析工程で取り出された音声出力内容を音声出力する音声出力工程と

ユーザからの音声入力を音声認識する音声認識工程と、

前記音声認識工程による認識結果を、前記文書解析工程で取り出された入力候補と照合して、当該認識結果に一致する入力候補に対応する指定情報に基づいて、前記文書取得工程による新たな文書の取得あるいは前記文書解析工程による次の解析を制御する制御工程とを備えることを特徴とする文書処理プログラムを格納した記憶媒体。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】

本発明は、所定のマークアップ言語で記述された文書を音声対話により処理する文書処理装置及びその方法、及びそのプログラムを格納した記憶媒体に関するものである。

【0 0 0 2】

【従来の技術】

Webのコンテンツへのアクセスは、従来、コンピュータのグラフィカルユーザインタフェースを用いたブラウザにより行われてきた。これに対し、最近、Webのコンテンツへのアクセスを、音声対話で行なうボイスブラウザが、電話を介したアクセスなどの目的で使われるようになってきた。

【0003】

ボイスブラウザでは、Webのコンテンツを音声で出力する。音声出力は、テキストで記述されたコンテンツを音声合成により音声に変換して出力する場合と、録音により音声データとして用意されたコンテンツを再生して出力する場合とがある。この音声出力は、グラフィカルユーザインタフェースでのブラウザにおけるページの表示に相当する。

【0004】

グラフィカルユーザインタフェースでのブラウザでは、次のコンテンツへの移動やフォームへの入力を、マウス操作やキーボード入力で行なうが、ボイスブラウザでは、これを音声入力で行なう。すなわち、ユーザの音声入力を音声認識して、認識結果を用いて、次のコンテンツへの移動やフォームへの入力を行なう。

【0005】

このようなボイスブラウザ用のコンテンツとして、専用のマークアップ言語を用いる方法がある。しかし、この方法では、そのコンテンツにグラフィカルユーザインタフェースのブラウザでアクセスすることができないし、現在多数存在するグラフィカルユーザインタフェース向けのコンテンツにアクセスすることもできない。そこで、グラフィカルユーザインタフェースのブラウザで用いられているマークアップ言語であるHTMLをボイスブラウザでも用いる方法がある。

【0006】

この方法では、HTMLで記述されたコンテンツから、特定のルールにより音声での出力内容と入力候補、すなわち音声認識語彙と入力に応じた処理内容を定める。例えば、次のようなルールを用いたボイスブラウザ装置がある。

【0007】

まず、出力内容は、ブラウジングの対象となっているHTML文書の先頭から最後

までのテキストとする。ただし、URLがHTML文書の途中を示す場合はその位置からとし、途中に<HR>タグがある場合はそのタグまでの間とする。入力候補は同じ範囲内のアンカー(<A>タグで囲まれる範囲のテキスト)とする。入力候補中に存在する語が入力された場合は、そのリンクのリンク先を新たなブラウジング対象として同様の処理を行なう。

【 0 0 0 8 】

例えば、図4に示すHTML文書を対象にした場合を考える。このHTML文書のURLは、“http://guide/index.html”であるとする。まずボイスブラウザは「店のジャンルを次から選んで下さい。フレンチ。イタリアン。」と音声出力し、ユーザの入力を待つ。ユーザが、例えば「イタリアン」と入力すると、“http://guide/index.html#italian”のHTML文書の位置から同様の処理を行なう。すなわち、「店を選んで下さい。▽▽。□□。」と音声出力し、ユーザの入力を待つ。ユーザが例えば、「▽▽」と入力すると、“http://guide/shop3.html”のHTML文書を取得して同様の処理を進める。

【 0 0 0 9 】

【発明が解決しようとする課題】

しかしながら、上記従来例の装置では、コンテンツを特定ルールに従って記述しなければならず、コンテンツをグラフィカルユーザインタフェース向けにも作成する場合の自由度が損なわれるという問題があった。

【 0 0 1 0 】

本発明は上記の問題に鑑みてなされたものであり、HTML等のグラフィカルユーザインタフェース用のマークアップ言語で記述されたコンテンツから、音声での出力内容と入力候補を定めるルールを複数用意し、そのルールのいずれを使用するかをコンテンツの作成者あるいはユーザが指定できるボイスブラウザ装置を提供することを目的とする。

【 0 0 1 1 】

【課題を解決するための手段】

上記の課題を解決するために、本発明の文書処理装置は以下のような構成を備える。即ち、指定された取得先より所定のマークアップ言語で記述された文書を

取得する文書取得手段と、音声入出力内容を規定するルールを所定の複数のルールより選択するルール選択手段と、前記ルール選択手段で選択されたルールに基づいて、前記文書取得手段で取得された文書の指定範囲を解析して、音声出力内容と、音声入力候補と、各音声入力候補に対応する次の処理対象を指定する指定情報とを取り出す文書解析手段と、前記文書解析手段で取り出された音声出力内容を音声出力する音声出力手段と、ユーザからの音声入力を音声認識する音声認識手段と、前記音声認識手段による認識結果を、前記文書解析手段で取り出された入力候補と照合して、当該認識結果に一致する入力候補に対応する指定情報に基づいて、前記文書取得手段による新たな文書の取得あるいは前記文書解析手段による次の解析を制御する制御手段とを備える。

【 0 0 1 2 】

また、本発明の他の態様によれば、文書処理方法に、指定された取得先より所定のマークアップ言語で記述された文書を取得する文書取得工程と、音声入出力内容を規定するルールを所定の複数のルールより選択するルール選択工程と、前記ルール選択工程で選択されたルールに基づいて、前記文書取得工程で取得された文書の指定範囲を解析して、音声出力内容と、音声入力候補と、各音声入力候補に対応する次の処理対象を指定する指定情報とを取り出す文書解析工程と、前記文書解析工程で取り出された音声出力内容を音声出力する音声出力工程と、ユーザからの音声入力を音声認識する音声認識工程と、前記音声認識工程による認識結果を、前記文書解析工程で取り出された入力候補と照合して、当該認識結果に一致する入力候補に対応する指定情報に基づいて、前記文書取得工程による新たな文書の取得あるいは前記文書解析工程による次の解析を制御する制御工程とを備える。

【 0 0 1 3 】

また、本発明の他の態様によれば、記憶媒体に、指定された取得先より所定のマークアップ言語で記述された文書を取得する文書取得工程と、音声入出力内容を規定するルールを所定の複数のルールより選択するルール選択工程と、前記ルール選択工程で選択されたルールに基づいて、前記文書取得工程で取得された文書の指定範囲を解析して、音声出力内容と、音声入力候補と、各音声入力候補に

対応する次の処理対象を指定する指定情報とを取り出す文書解析工程と、前記文書解析工程で取り出された音声出力内容を音声出力する音声出力工程と、ユーザからの音声入力を音声認識する音声認識工程と、前記音声認識工程による認識結果を、前記文書解析工程で取り出された入力候補と照合して、当該認識結果に一致する入力候補に対応する指定情報に基づいて、前記文書取得工程による新たな文書の取得あるいは前記文書解析工程による次の解析を制御する制御工程とを備える文書処理プログラムを格納する。

【0014】

【発明の実施の形態】

以下、図面を参照して本発明の一実施形態を詳細に説明する。

【0015】

〔実施形態1〕

図1は、実施形態1に係るボイスブラウザ装置の基本構成を示すブロック図である。

【0016】

同図において、101は、指定のHTML文書を取得するHTML文書取得部である。102は、HTML文書取得部101で取得されたHTML文書を保持するHTML文書保持部である。103は、HTML文書保持部102に保持されたHTML文書中に指定された、音声入出力内容を定めるルールを取得する指定ルール取得部である。104は、指定ルール取得部103で取得された指定ルールを保持する指定ルール保持部である。

【0017】

105は、指定ルール保持部104に保持されたルールに従って、HTML文書保持部102に保持されたHTML文書を解析して、音声入出力の内容を取り出すHTML文書解析部である。106は、HTML文書解析部105で解析され取り出された音声入出力内容を保持する入出力内容保持部である。107は、入出力内容保持部106に保持された音声出力内容を必要に応じて音声合成して音声出力する音声出力部である。

【0018】

108は、ユーザからの音声入力を受け入れて音声認識する音声入力部である。109は、音声入力部108における入力内容の認識結果を、入出力内容保持部106に保

持されている入力候補と照らし合わせて、HTML文書取得部101による新しいHTML文書の取得や、HTML文書解析部105によるHTML文書の解析を制御するブラウザ制御部である。

【0019】

図2は、本実施形態のボイスブラウザ装置のハードウェア構成を示す図である。同図において、201はCPUであり、後述する手順を実現するプログラムに従って動作する。202はRAMであり、HTML文書保持部102と指定ルール保持部104と入出力内容保持部106と上記プログラムの動作に必要な記憶領域とを提供する。203はディスク装置であり、後述する手順を実現するプログラムを保持する。

【0020】

204はスピーカーであり、音声出力部107で生成する音声データを出力する。205はマイクロホンであり、音声入力部108で処理する音声データを入力する。206はネットワークインタフェースであり、HTML文書取得部101がネットワーク経由でHTML文書を取得する際のネットワークを介した通信を実現する。207は、上記各部を接続するバスである。

【0021】

以下、図3のフローチャートを参照して、本実施形態のボイスブラウザ装置の処理手順を説明する。

【0022】

まず、ステップS301では、初期HTML文書を取得する。初期のHTML文書は、ボイスブラウザ装置で予め定められた文書、ユーザが設定した文書、前回最後にアクセスした文書など任意の文書を取ることができる。いずれも初期HTML文書のURLを定めておくことになる。HTML文書取得部101がこのURLに応じてネットワークを介してHTTPリクエストを送出して取得するか、装置内のディスク装置に既に保持されたファイルから取得する。取得したHTML文書はHTML文書保持部102に保持する。そしてステップS302に移る。

【0023】

ステップS302では、HTML文書保持部102に保持されているHTML文書から、その文書中に記述されている、音声入出力内容を定めるルールを指定するデータを取

得する。本実施形態では、HTML文書中の<VB>タグの属性MODEの値によってルールが指定される。この値を指定ルール保持部104に保持する。値に応じたルールそのものは、HTML文書解析部105のプログラムとして予め装置に組み込まれている。例えば、図5に示すHTML文書进行处理する場合は、Hを指定ルール保持部104に保持する。HTML文書中に<VB>タグがない場合は、Lを指定ルール保持部104に保持する。そしてステップS303に移る。

【 0 0 2 4 】

ここで、本実施形態で使用するルールを説明する。本実施形態で、指定ルール保持部104の値がHの場合のルールは、以下の通りである。最初の出力内容は、<VB>タグのOUTPUT属性の値と、次に説明する入力候補とする。入力候補は、HTML文書中の<H>タグで囲まれるそれぞれの見出しとする。入力候補に含まれる語句が入力された場合は、以下の処理を行なう。まず、次の出力内容を、選択された見出しから次の<H>タグあるいは文書の最後までテキストとする。そして、入力候補は同じ範囲内のアンカーとする。入力候補に含まれる語句が入力された場合は、そのリンクのリンク先を新たなブラウジング対象として同様の処理を行なう。

【 0 0 2 5 】

一方、本実施形態で、指定ルール保持部104の値がLの場合のルールは、従来技術として説明した処理手順を行なうというルールである。すなわち、出力内容を、ブラウジングの対象となっているHTML文書の先頭から最後までテキストとする。ただし、URLがHTML文書の途中を示している場合はその位置からとし、途中に<HR>タグがある場合はそのタグまでの間とする。また、入力候補は同じ範囲内のアンカー(<A>タグで囲まれる範囲のテキスト)とする。入力候補に含まれる語句が入力された場合は、そのリンクのリンク先を新たなブラウジング対象として同様の処理を行なう。

【 0 0 2 6 】

ステップS303では、指定ルール保持部104に保持されている値に応じたルールに従って、HTML文書保持部102に保持されているHTML文書を解析して、音声入力の内容を取り出し、入出力内容保持部106に保持する。そしてステップS304に

移る。

【 0 0 2 7 】

図6は、本実施形態における入出力内容保持部106の例を示す図である。601は、音声出力内容となるテキストを保持する領域である。602は、入力候補とそれぞれの処理を定めるデータを保持する領域である。図6では、入力候補ごとのデータを1つの行で示している。各行において、列603には、入力候補が保持される。列604には、その候補が入力された後の処理対象となるHTMLページの示すURLが保持される。列605には、指定ルールがHの場合に、次に処理対象とする先頭の見出しのパターンが保持される。

【 0 0 2 8 】

ステップS303では、指定ルール保持部104に保持されている値がHの場合は、ステップS302から移ってきた場合と、ステップS307から移ってきた場合とで処理が異なる。

【 0 0 2 9 】

前者では、<VB>タグのOUTPUT属性の値と、次に説明する入力候補とを入出力内容保持部106の領域601に保持する。HTML文書中の<H>タグで囲まれるそれぞれの見出しを、入力候補として列603に保持する。そして、それぞれの候補に対して、現在処理中のHTML文書のURLを列604に保持する。また、各見出しのタグを含むパターンを列605に保持する。

【 0 0 3 0 】

後者では、ステップS306で選択された入力候補の列605のパターンを、HTML文書保持部102に保持されているHTML文書から探し、そこから次の<H>タグあるいは文書の最後までテキストを入出力内容保持部106の領域601に保持する。そして、同じ範囲にあるアンカーを入力候補として、それぞれの候補に対して、そのリンク先のURLを列604に保持する。列605は空とする。

【 0 0 3 1 】

指定ルール保持部104に保持されている値がLの場合は、HTML文書の先頭から最後までテキストとする。ただし、URLがHTML文書の途中を示す場合はその位置からとし、途中に<HR>タグがある場合はタグまでの間とする。そして、入力候補

は同じ範囲内のアンカーとして、それぞれの候補に対して、そのリンク先のURLを列604に保持する。列605は空とする。

【 0 0 3 2 】

図6は、図5に示すHTML文書进行处理した時の入出力内容保持部106の状態である。

【 0 0 3 3 】

ステップS304では、入出力内容保持部106の領域601に出力内容として保持されたテキストを音声合成して音声データに変換し、スピーカーへ出力する。そしてステップS305に移る。

【 0 0 3 4 】

ステップS305では、マイクロホンへ一定レベル以上の音声入力がある一定時間以上続いた場合にその音声データを音声認識する。音声認識に成功すればステップS306に移る。音声入力がない場合や認識に失敗した場合はステップS305を繰り返す。

【 0 0 3 5 】

ステップS306では、ステップS305の音声認識結果を入出力内容保持部106に保持されている入力候補と比較する。一致する入力候補がある場合はステップS307に移る。一致する入力候補がない場合はステップS305に戻る。

【 0 0 3 6 】

ステップS307では、ステップS306で選択された入力候補に列605の見出しパターンのデータがあるかを調べ、データがある場合はステップS303に戻り、当該データ以後を取り出す処理を行なう。ない場合はステップS308に移る。

【 0 0 3 7 】

ステップS308では、ステップS306で一致した入力候補のURLで示されるHTML文書を新たに取得し、HTML文書保持部102に保持する。そして、ステップS302に戻る。

【 0 0 3 8 】

HTML文書保持部102に図5のHTML文書が保持されており、入出力内容保持部106の状態が図6である時に、「イタリアン」が入力されると、入出力内容保持部106

の状態は新たに図7のようになる。従って、HTML文書保持部102に図5のHTML文書が保持されてからの入出力は次のようになる。

出力:店のジャンルを選び、次に店を選んで下さい。フレンチ。イタリアン。

入力:イタリアン

出力:イタリアン。▽▽。□□。

入力:▽▽

出力:……

【 0 0 3 9 】

HTML文書保持部102に、図4のHTML文書が最初に保持された場合の入出力は、次のようになる。

出力:店のジャンルを次から選んで下さい。フレンチ。イタリアン。

入力:イタリアン

出力:イタリアン。店を選んで下さい。▽▽。□□。

入力:▽▽

出力:……

【 0 0 4 0 】

図4のHTML文書を通常のブラウザで表示した例が図8であり、図5のHTML文書を通常のブラウザで表示した例が図9である。このように、本実施形態のボイスブラウザ装置を利用すると、同様の音声対話を実現するコンテンツを、違う形式で表示するような複数の記述が可能になる。

【 0 0 4 1 】

一方、HTML文書保持部102に図10のHTML文書が最初に保持された場合の入出力は次のようになる。

出力:店を次から選んで下さい。フレンチ。○○。△△。イタリアン。▽▽。□□。

入力:▽▽

出力:……

【 0 0 4 2 】

図10のHTML文書は、図5のHTML文書とは、<VB>タグのMODE属性の値が異なるだ

けである。本実施形態のボイスブラウザ装置を用いれば、一部のタグを変えるだけで、同様のHTML文書に対する音声対話の内容を変えることができる。

【0043】

〔実施形態2〕

上記実施形態1においては、入出力内容を決定するルールをコンテンツ中で指定する場合について説明したが、これに限定されるものではなく、ユーザが指定してもよい。また、コンテンツの指定とユーザの指定を両方受け入れ可能とし、そのいずれかを優先させてもよい。

【0044】

図11は、実施形態2に係る装置の基本構成を示すブロック図である。同図において、101～103、105～109は、図1と同様である。以下、図1と相違する部分について説明する。

【0045】

1101は、ユーザが定めるルールを保持するユーザルール保持部である。1102は、指定ルール取得部103で取得された指定ルールとユーザルール保持部1101に保持されているユーザルールのどちらを優先するかを決定する解析ルール決定部である。1103は、解析ルール決定部1102で決定された解析ルールを保持する解析ルール保持部である。そして、HTML文書解析部105では、解析ルール保持部1103に保持されたルールに従って、HTML文書保持部102に保持されたHTML文書を解析して、音声入出力の内容を取り出す。

【0046】

本実施形態では、コンテンツの指定ルールとユーザルールのどちらを優先するかが問題になるが、例えば、どちらかを常に優先するようにすればよい。また、どちらを優先するかを、ユーザに選択させることもできる。あるいは、HTML文書にルールを指定するタグがない時はユーザルールを使用し、タグがある場合はHTML文書で指定するルールを優先してもよい。

【0047】

〔他の実施形態〕

上記実施形態においては、ユーザルールはHTML文書に関係なく同一である場合

について説明したが、これに限定されるものではなく、HTML文書ごとに変えてもよい。HTML文書进行处理して音声出力した後、特定のコマンド(例えば「リストモード」)を入力すると、ユーザルール保持部に保持されるユーザルールを変更してもよい。

【 0 0 4 8 】

上記実施形態においては、ユーザルールを変更した時に、変更結果が次の対話から反映される場合について説明したが、これに限定されるものではなく、変更した時の対象HTML文書から反映させてもよい。それには、ユーザルール保持部の内容が変更されたら、HTML文書の解析処理からを再度行なえばよい。

【 0 0 4 9 】

上記実施形態においては、ユーザが直接指定するルールをユーザルールとする場合について説明したが、これに限定されるものではなく、HTML文書ごとに適用するルールを保持しておいて、そのHTML文書进行处理するごとに保持していたルールを適用してもよい。これは、HTML文書のURLと適用すべきルールとを対応付けたテーブルを保持しておき、HTML文書を取得するたびにURLを用いてテーブルを検索し、テーブルに当該URLが保持されている場合は、対応するルールをユーザルール保持部に保持させることで実現できる。

【 0 0 5 0 】

例えば、図5のHTML文書に対して、ジャンルで絞ってから店を選択するか、直接店を選択するかを、ユーザが設定しておくことが可能になる。

【 0 0 5 1 】

上記実施形態においては、音声の入出力を装置に直接接続されているスピーカとマイクで行なう場合について説明したが、これに限定されるものではなく、他の音声入出力デバイスを使用してもよい。V例えば、装置と電話回線を介して接続される電話機を使用してもよい。

【 0 0 5 2 】

図12は、電話機を使用する場合の本実施形態の情報提示装置のハードウェア構成を示す図である。同図において、201～203、206～207は、図2と同様である。1201は電話回線インタフェースであり、音声出力部107で生成された音声データ

を電話回線を介して外部の電話機に送信し、音声入力部108で処理すべき音声データを電話回線を介して外部の電話機から受信する。

【0053】

上記実施形態においては、ボイスブラウザ装置の入出力を全て音声で行なう場合について説明したが、これに限定されるものではなく、一部に音声以外の入力手段を用いてもよい。例えば、入力候補を音声入力するのではなく、候補の番号をキー入力してもよい。

【0054】

例えば、上述した電話機を使用する構成の場合は、電話機のダイヤルボタンから番号を入力し、そのトーンを受信することで番号入力を受け付けることができる。番号の付与の仕方としては、例えば、HTML文書での入力候補の出現順に1から昇順に付与する方法がある。また、このために、入力候補と共に番号を出力してもよい。例えば、入出力内容保持部106の状態が図6である時に、「店のジャンルを選び、次に店を選んで下さい。1番、フレンチ。2番、イタリアン。」と出力してもよい。

【0055】

上記実施形態においては、音声入出力内容を定めるルールとして、見出しを読み上げるルールと、<HR>タグまでのテキストを読み上げるルールとの二つを切り替える場合について説明したが、これに限定されるものではなく、様々なルールを指定してよい。例えば、上述した入力候補の番号を出力するかどうかを定めるルールを指定してもよい。HTML文書での指定方法として、<VB>タグにNUMBER属性を設けて、値がONのときは出力し、値がOFFの時は出力しないという方法を用いることができる。

【0056】

上記実施形態においては、HTML文書でルールを指定する方法として、<VB>タグを用いる場合について説明したが、これに限定されるものではなく、他のタグを使用してもよい。また、例えば<BODY>タグの属性に加えてもよい。あるいは、コメントに埋め込んでもよい。

【0057】

上記実施形態においては、ルールをボイスブラウザ装置に予め組み込んでおき、そのルールに対応するラベルを指定する場合について説明したが、これに限定されるものではなく、ルール自体を外部から指定できるようにしてもよい。例えば、上記実施形態では出力の対象は全てのテキストであったが、出力する内容を特定のタグで囲まれる部分だけとし、そのタグをHTML文書中で列挙してもよい。例えば、<VB>タグのOUTTAG属性の値として列挙すればよい。あるいは、上記実施形態では出力の終点とするタグが<HR>タグに固定されていたが、出力の終点とするタグをHTML文書中で指定してもよい。例えば、<VB>タグのENDTAG属性の値として指定すればよい。

【 0 0 5 8 】

上記実施形態においては、HTML文書を対象としたが、これに限定されるものではなく、HTMLを一部拡張・変更したマークアップ言語や他のマークアップ言語で記述された文書を対象としてもよい。

【 0 0 5 9 】

上記実施形態においては、音声認識の認識語彙を予め用意しておく場合について説明したが、これに限定されるものではなく、入力候補から作成してもよい。

【 0 0 6 0 】

上記実施形態においては、音声出力が終わってから音声入力を受け付ける場合について説明したが、これに限定されるものではなく、音声出力の途中で音声入力を受け付けてもよい。

【 0 0 6 1 】

上記実施形態においては、動作に必要なプログラムをディスク装置に保存する場合について説明したが、これに限定されるものではなく、任意の記憶媒体を用いて実現してもよい。また、同様の動作をする回路で実現してもよい。

【 0 0 6 2 】

なお、本発明は、複数の機器から構成されるシステムに適用しても、1つの機器からなる装置に適用してもよい。前述した実施形態の機能を実現するソフトウェアのプログラムコードを記録した記録媒体を、システム或いは装置に供給し、そのシステム或いは装置のコンピュータ(またはCPUやMPU)が記録媒体に格納され

たプログラムコードを読み出し実行することによっても、達成されることは言うまでもない。

【0063】

この場合、記録媒体から読み出されたプログラムコード自体が前述した実施形態の機能を実現することになり、そのプログラムコードを記録した記録媒体は本発明を構成することになる。

【0064】

プログラムコードを供給するための記録媒体としては、例えば、フロッピーディスク、ハードディスク、光ディスク、光磁気ディスク、CD-ROM、CD-R、磁気テープ、不揮発のメモリカード、ROMなどを用いることができる。

【0065】

また、コンピュータが読み出したプログラムコードを実行することにより、前述した実施形態の機能が実現されるだけでなく、そのプログラムコードの指示に基づき、コンピュータ上で稼働しているOSなどが実際の処理の一部または全部を行ない、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0066】

更に、記録媒体から読み出されたプログラムコードが、コンピュータに挿入された機能拡張ボードやコンピュータに接続された機能拡張ユニットに備わるメモリに書き込まれた後、そのプログラムコードの指示に基づき、その機能拡張ボードや機能拡張ユニットに備わるCPUなどが実際の処理の一部または全部を行ない、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0067】

【発明の効果】

以上説明したように、本発明によれば、所定のマークアップ言語で記述された文書から、音声での出力内容と入力候補とを定めるルールを複数用意し、そのルールのいずれを使用するかを文書の作成者あるいはユーザが指定できるようにしたので、文書の内容部分を変えずに、音声入出力内容を容易に変更できるという

効果が得られる。

【図面の簡単な説明】

【図 1】

実施形態 1 に係るボイスブラウザ装置の基本構成を示すブロック図である。

【図 2】

実施形態 1 に係るボイスブラウザ装置のハードウェア構成を示す図である。

【図 3】

実施形態 1 における処理の概要を示すフローチャートである。

【図 4】

従来装置が扱う HTML 文書の例を示す図である。

【図 5】

実施形態の装置が扱う HTML 文書の例を示す図である。

【図 6】

入出力内容保持部の具体例を示す図である。

【図 7】

入出力内容保持部の具体例を示す図である。

【図 8】

従来装置が扱う HTML 文書を表示した例を示す図である。

【図 9】

従来装置が扱う別の HTML 文書を表示した例を示す図である。

【図 10】

実施形態の装置が扱う別の HTML 文書の例を示す図である。

【図 11】

実施形態 2 の基本構成を示す図である。

【図 12】

他の実施形態のハードウェア構成を示す図である。

【符号の説明】

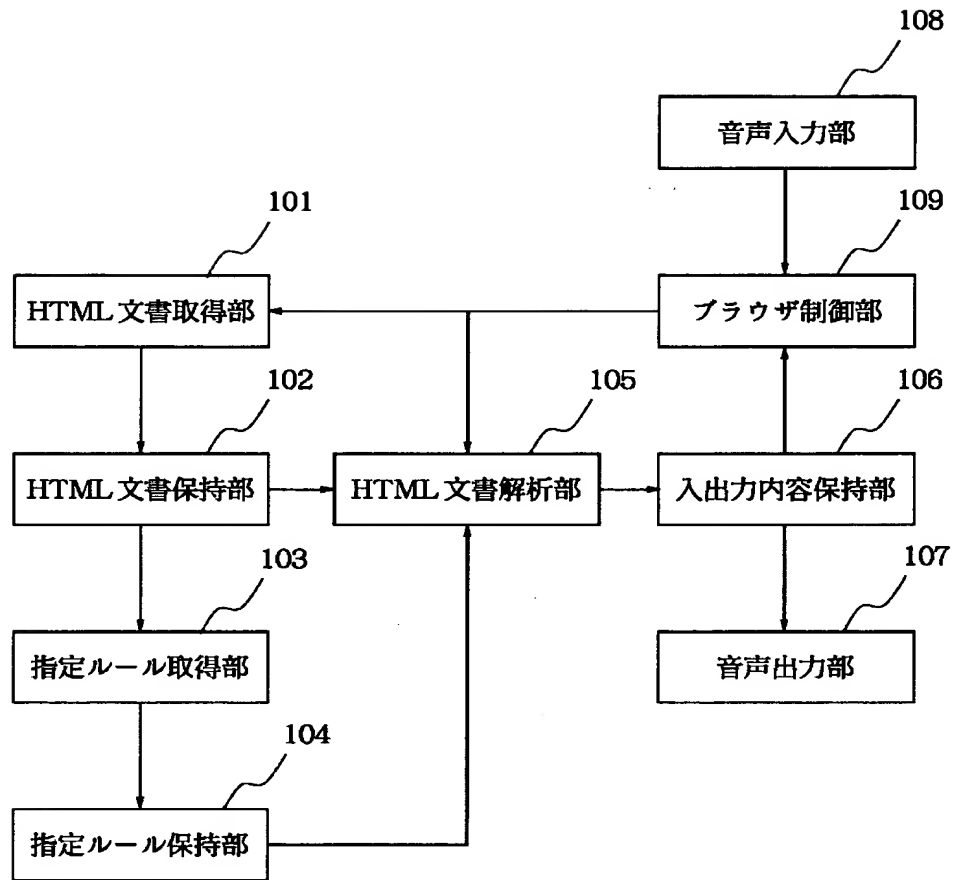
101 HTML 文書取得部

102 HTML 文書保持部

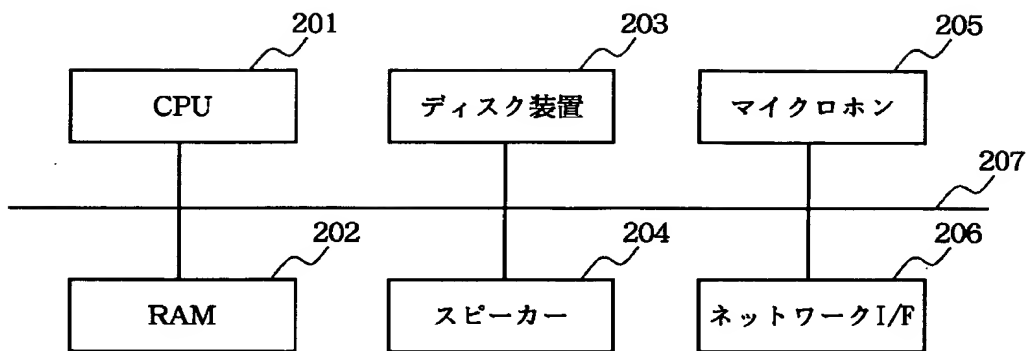
- 1 0 3 指定ルール取得部
- 1 0 4 指定ルール保持部
- 1 0 5 HTML文書解析部
- 1 0 6 入出力内容保持部
- 1 0 7 音声出力部
- 1 0 8 音声入力部
- 1 0 9 ブラウザ制御部
- 2 0 1 CPU
- 2 0 2 RAM
- 2 0 3 ディスク装置
- 2 0 4 スピーカー
- 2 0 5 マイクロホン
- 2 0 7 バス
- 6 0 1 音声出力するテキストを保持する領域
- 6 0 2 入力候補と次の処理に関するデータを保持する領域
- 6 0 3 各行において入力候補を保持する領域
- 6 0 4 各行において次に処理するHTML文書のURLを保持する領域
- 6 0 5 各行において次に処理を開始する見出しのパターンを保持する領域
- 1 1 0 1 ユーザルール保持部
- 1 1 0 2 解析ルール決定部
- 1 1 0 3 解析ルール保持部
- 1 2 0 1 電話インタフェース

【書類名】 図面

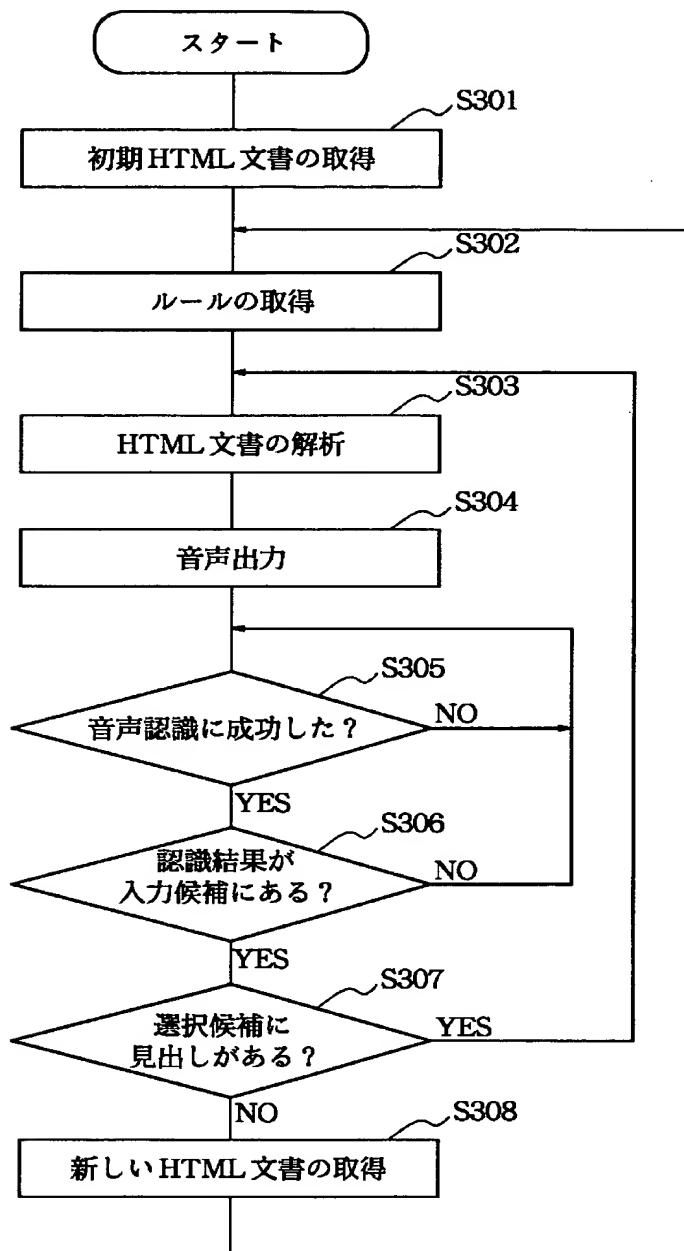
【図 1】



【図 2】



【図 3】



【図 4】

```
<!DOCTYPE HTML PUBLIC "-//IETF//DTD HTML//EN">
<HTML>
<HEAD>
<TITLE>Restaurant Guide</TITLE>
</HEAD>
<BODY>
店のジャンルを次から選んで下さい。
<UL>
<LI><A HREF="#french">フレンチ</A>
<LI><A HREF="#italian">イタリアン</A>
</UL>
<HR>
<A NAME="french">フレンチ</A><BR>
店を選んで下さい。
<UL>
<LI><A HREF="shop1.html">〇〇</A>
<LI><A HREF="shop2.html">△△</A>
</UL>
<HR>
<A NAME="italian">イタリアン</A><BR>
店を選んで下さい。
<UL>
<LI><A HREF="shop3.html">▽▽</A>
<LI><A HREF="shop4.html">□□</A>
</UL>
</BODY>
</HTML>
```

【図 5】

```
<!DOCTYPE HTML PUBLIC "-//IETF//DTD HTML//EN">
<HTML>
<HEAD>
<TITLE>Restaurant Guide</TITLE>
</HEAD>
<BODY>
<VB MODE="H" output="店のジャンルを選び,次に店を選んで下さい.">
店を選んで下さい.
<H2>フレンチ</H2>
<UL>
<LI><A HREF="shop1.html">〇〇</A>
<LI><A HREF="shop2.html">△△</A>
</UL>
<H2>イタリアン</H2>
<UL>
<LI><A HREF="shop3.html">▽▽</A>
<LI><A HREF="shop4.html">□□</A>
</UL>
</BODY>
</HTML>
```

【図 6】

601

店のジャンルを選び、次に店を選んで下さい。フレンチ.イタリアン

602

603 フレンチ	604 http://guide/index2.html	605 <H2>フレンチ</H2>
イタリアン	http://guide/index2.html	<H2>イタリアン</H2>

【図 7】

イタリアン.▽▽.□□.		
▽▽	http://guide/shop3.html	
□□	http://guide/shop4.html	

【図 8】

<p>店のジャンルを次から選んで下さい。</p> <ul style="list-style-type: none">・ フレンチ・ イタリアン <hr/> <p>フレンチ 店を選んで下さい。</p> <ul style="list-style-type: none">・ ○○・ △△ <hr/> <p>イタリアン 店を選んで下さい。</p> <ul style="list-style-type: none">・ ▽▽・ □□
--

【図 9】

店を選んで下さい。

- ・フレンチ
- ・○○
- ・△△

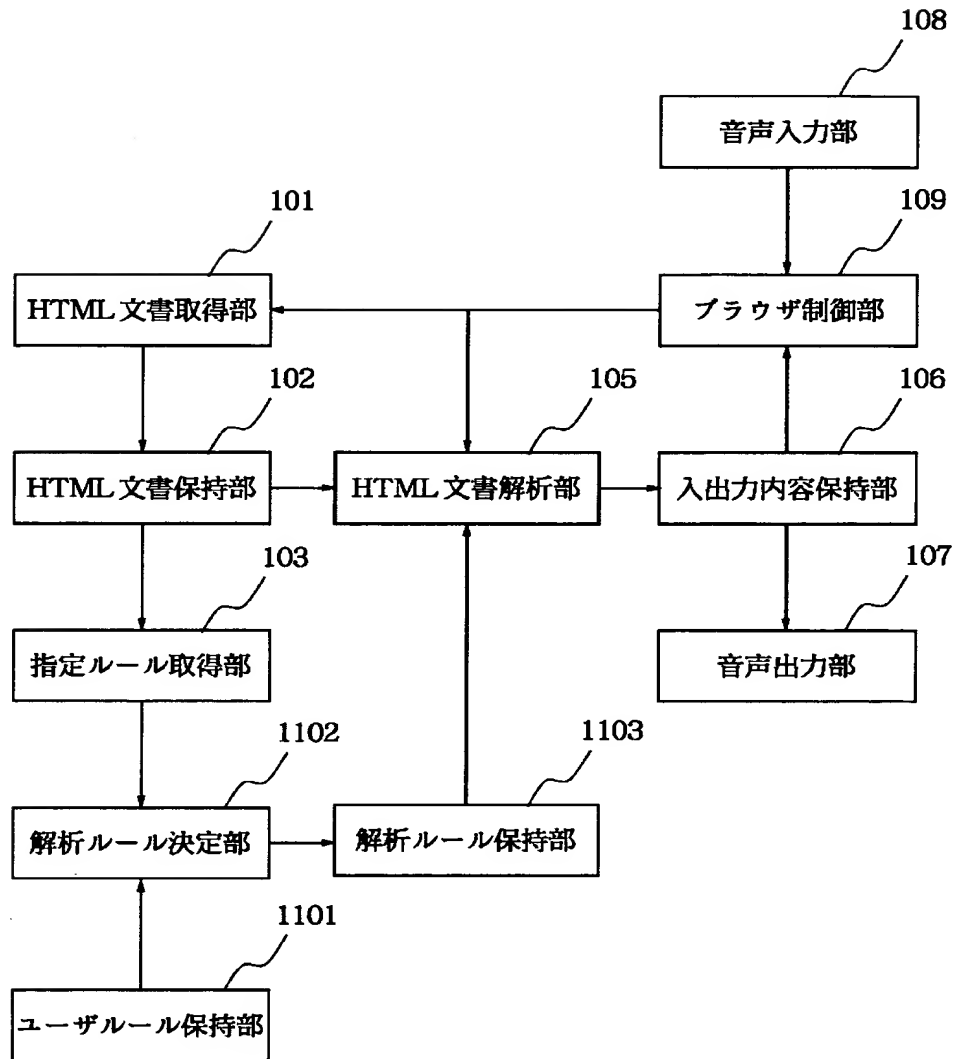
イタリアン

- ・▽▽
- ・□□

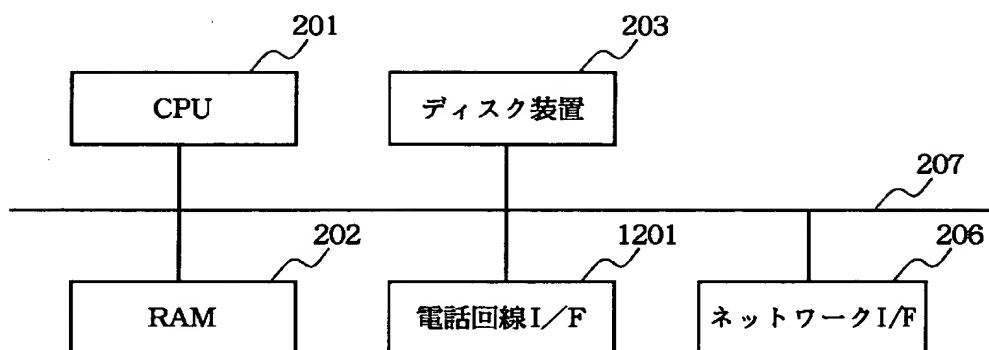
【図 1 0】

```
<!DOCTYPE HTML PUBLIC "-//IETF//DTD HTML//EN">
<HTML>
<HEAD>
<TITLE> Restaurant Guide</TITLE>
</HEAD>
<BODY>
<VB MODE="L">
店を選んで下さい。
<H2> フレンチ</H2>
<UL>
<LI><A HREF="shop1.html">〇〇</A>
<LI><A HREF="shop2.html">△△</A>
</UL>
<H2> イタリアン</H2>
<UL>
<LI><A HREF="shop3.html">▽▽</A>
<LI><A HREF="shop4.html">□□</A>
</UL>
</BODY>
</HTML>
```

【図 1 1】



【図 1 2】



【書類名】 要約書

【要約】

【課題】 グラフィカルユーザインタフェース向けのコンテンツを、音声入出力により簡単にブラウジングできるようにする。

【解決手段】 ボイスブラウザ装置に、指定されたURLよりHTML文書を取得するHTML文書取得部101と、音声入出力内容を規定するルールを取得する指定ルール取得部103と、そのルールに基づいて、取得文書の指定範囲を解析して、音声出力内容と、音声入力候補と、各音声入力候補に対応する次の処理対象を指定する指定情報とを取り出すHTML文書解析部105と、取り出された音声出力内容を音声出力する音声出力部107と、ユーザからの音声入力を音声認識する音声入力部108と、認識結果と一致する入力候補に対応する指定情報に基づいて、新たな文書の取得あるいは次の解析を制御するブラウザ制御部109とを備える。

【選択図】 図 1

出 願 人 履 歴 情 報

識別番号 [000001007]

1. 変更年月日	1990年 8月30日
[変更理由]	新規登録
住 所	東京都大田区下丸子3丁目30番2号
氏 名	キャノン株式会社